# Fruits, vegetables, and hMLH1 protein-deficient and -proficient colon cancer: The Netherlands cohort study

# Fruits, Vegetables, and hMLH1 Protein–Deficient and –Proficient Colon Cancer: The Netherlands Cohort Study

Petra A. Wark,[1] Matty P. Weijenberg,[2] Pieter van 't Veer,[1] Gerda van Wijhe,[1] Margreet Lüchtenborg,[2] Goos N.P. van Muijen,[4] Anton F.P.M. de Goeij,[3] R. Alexandra Goldbohm,[5] and Piet A. van den Brandt[2]

[1]Division of Human Nutrition, Wageningen University, Wageningen, the Netherlands; [2]Nutrition and Toxicology Research Institute Maastricht (NUTRIM), Department of Epidemiology; [3]Research Institute Growth and Development (GROW), Department of Pathology, Maastricht University, the Netherlands; [4]Department of Pathology, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands; and [5]TNO Nutrition and Food Research, Zeist, the Netherlands

## Abstract

Background: **Clinical and pathologic differences exist between colon carcinomas deficient and proficient in the mismatch repair protein hMLH1. Animal and *in vitro* studies suggest that fruits, vegetables, folate, and antioxidants are associated with colonic expression of mismatch repair genes.**

Methods: **Associations between consumption of fruits and vegetables and hMLH1 protein–deficient and –proficient colon cancer were evaluated in the Netherlands Cohort Study on diet and cancer using a case-cohort approach. A self-administered food frequency questionnaire was completed, in 1986, by 120,852 individuals ages 55 to 69 years. Using immunohistochemistry, hMLH1 protein expression was assessed in colon cancer tissue obtained from 441 patients who were identified over 7.3 years of follow-up excluding the initial 2.3 years. Incidence rate ratios (RR) were estimated for hMLH1 protein–deficient and –proficient colon cancer.**

Results: **hMLH1 protein expression was absent in 54 tumors (12.2%) and present in 387 tumors. Fruit consumption was associated with hMLH1 protein–deficient colon cancer [highest versus lowest tertile, RR, 0.46; 95% confidence interval (95% CI), 0.23-0.90; $P_{trend}$ = 0.029] but not with hMLH1 protein–proficient tumors (highest versus lowest tertile, RR, 1.03; 95% CI, 0.78-1.35; $P_{trend}$ = 0.81). Total consumption of vegetables was not associated with either type of tumor (hMLH1 protein deficient: RR, 0.86; 95% CI, 0.45-1.65; $P_{trend}$ = 0.67; hMLH1 protein proficient: RR, 0.94; 95% CI, 0.72-1.23; $P_{trend}$ = 0.72). No associations were observed for folate, fiber, antioxidants, or subgroups of vegetables.**

Conclusion: **These analyses indicate that an inverse association between consumption of fruits and colon cancer may be confined to the subgroup of tumors with a deficient mismatch repair system.** (Cancer Epidemiol Biomarkers Prev 2005;14(7):1619–25)

## Introduction

Between 80% and 90% of colorectal carcinomas are thought to arise via the chromosome instability pathway, which is associated with mutations in adenomatous polyposis coli (*APC*), Kirsten-ras (*K-ras*), and *p53* genes. About 10% to 20% of colon carcinomas appear to be arising via the hypermutability pathway involving deficiency of DNA mismatch repair (MMR) enzymes. This pathway is associated with microsatellite instability (MSI) and mutations in genes such as insulin-like growth factor II receptor (*IGFIIR*) and Bcl-2-associated X protein (*BAX*; ref. 1). In sporadic colorectal cancer, *hMLH1* gene inactivation by promoter hypermethylation is thought to be the main mechanism behind MMR deficiency (2-8).

Carcinomas showing MMR deficiency have clinical and pathologic features that distinguish them from carcinomas with MMR proficiency. More often located in the proximal colon and showing a mucinous or undifferentiated histology and prominent lymphocytic tumor infiltration, MMR-deficient carcinomas have a relatively favorable prognosis and respond differently to chemotherapeutic agents compared with MMR-proficient carcinomas (9-13). Given these distinct tumor features, different sets of dietary and lifestyle factors may be involved in the etiology of MMR-deficient and -proficient colorectal carcinomas.

Fruits and vegetables are candidates for such potential differential effects. Folate, which is present in, among others, green leafy vegetables and citrus fruits, can affect *S*-adenosylmethionine levels, which regulate DNA methylation (14). Antioxidants could have an influence because MMR enzymes are involved in repair of oxidative DNA damage, and MSI was reduced in MMR-deficient colon cells growing in the presence of ascorbate (15). Moreover, in a case-control study, fruit consumption was inversely associated with MSI-high colon carcinomas with hypermethylated hMLH1 but was not associated with MSI-high colon carcinomas without hypermethylated hMLH1 (16). In another case-control study, consumption of vegetables was inversely associated with colon carcinomas without MSI (17). However, an inverse association between consumption of vegetables and colon carcinomas with MSI was also observed, although not statistically significant (17).

In the Netherlands Cohort Study on diet and cancer, we examined whether consumption of fruits and vegetables is differentially associated with hMLH1 protein–deficient and –proficient colon cancer.

## Materials and Methods

**Cohort.** The Netherlands Cohort Study on diet and cancer is a population-based prospective study that was initiated in

Received 2/10/05; revised 4/15/05; accepted 4/25/05.

**Grant support:** The Netherlands Organisation for Scientific Research and the Dutch Cancer Society.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

**Requests for reprints:** Matty P. Weijenberg, Department of Epidemiology, Maastricht University, P.O. Box 616, 6200 MD Maastricht, the Netherlands.
E-mail: mp.weijenberg@epid.unimaas.nl

September 1986. At that time, 58,279 men and 62,573 women ages 55 to 69 years completed a mailed, self-administered questionnaire on dietary habits and other potential risk factors for cancer. After baseline exposure assessment, a subcohort of 3,500 people (1,688 men and 1,812 women) was randomly selected and their vital statuses were followed up biennially to estimate accumulated person-years (18).

**Follow-up for Cancer.** Cancer follow-up consists of annual linkage of the entire study database to the Netherlands Cancer Registry and the nationwide registry of cytopathology and histopathology (PALGA). The completeness of the follow-up procedure was estimated to be nearly 100% (19, 20). No subcohort members were lost to follow-up.

The initial 2.3 years of follow-up were excluded because of incomplete coverage of PALGA and to exclude misclassification of exposure related to potential subclinical disease. Subcohort members with prevalent cancer other than nonmelanoma skin cancer were also excluded, resulting in a subcohort size of 3,264 individuals.

Between 1989 and 1994, 929 incident cases with histologically confirmed colorectal carcinomas were identified, 819 of which (88%) could be identified in the PALGA database; data from PALGA were needed to identify in which pathologic laboratory tumor tissue was stored. Tumor characteristics (Dukes' stage, sublocalization, and differentiation grade) were obtained from the database of the Netherlands Cancer Registry. We classified cancers as proximal colon cancer (*International Classification of Diseases for Oncology, first edition* codes 153.0, 153.1, 153.4, 153.5, and 153.6), distal colon cancer (codes 153.2, 153.3, and 153.7), rectosigmoid (code 154.0), and rectal cancer (code 154.1).

**Tumor Tissue Samples.** After approval by the Ethical Review Boards of Maastricht University, PALGA, and the Netherlands Cancer Registry, we were able to retrieve formalin-fixed, paraffin-embedded colorectal tumor tissue from 776 of the 819 patients (95%) from 54 pathology laboratories throughout the country. Thirty-nine tumor blocks could not be used: 20 contained only normal mucosa, 10 turned out to be adenomas, three were noncolorectal malignancies, and tissue was sparse in six. The remaining 737 (95%) specimens were distributed as follows: proximal colon ($n$ = 240), distal colon ($n$ = 224), colon cancer not otherwise specified ($n$ = 12), rectosigmoid ($n$ = 85), and rectum ($n$ = 176).

**hMLH1 Immunohistochemical Analyses.** Immunohistochemical analyses were done on 4-μm sections of formalin-fixed, paraffin-embedded colorectal cancer tissue and adjacent normal tissue using monoclonal antibody against hMLH1 (clone G168-15; dilution 1:100; BD PharMingen International/Becton Dickinson, San Diego, CA) as previously described (16). Absence of hMLH1 protein expression was scored upon the absence of any nuclear staining in tumor cells in combination with a signal in positive internal control tissue (normal epithelial cells, stromal cells, muscle cells, and lymphocytes). Expression of hMLH1 was considered present in case of a nuclear signal in tumor cells and positive internal control tissue. Two investigators (P.A.W. and G. van W.) reviewed the immunostaining independently. Discrepancies were reexamined and discussed with a pathologist until consensus was reached.

hMLH1 expression status was determined successfully in 725 (98%) of the 737 samples: 234 occurred in the proximal colon, 222 in the distal colon, 12 were colon cancers not otherwise specified, 84 occurred in the rectosigmoid, and 173 in the rectum.

**Restriction to Colon Cancer.** As colon and rectum cancer are considered two distinct disease entities (21, 22), have a different etiology (23), and the rectosigmoid can be considered a clinically applied term rather than an anatomically defined transitional zone between the colon and rectum, it would be ideal to stratify analyses to site (colon, rectum, and rectosigmoid). Because hMLH1 protein expression was detected in cancer tissue specimens of all 84 patients with rectosigmoid cancer and in 169 of the 173 rectal cancer specimens, such analyses were not feasible. Therefore, we restricted the present analyses to the 468 colon cancer patients.

**Exposure Assessment.** At baseline, a 150-item semiquantitative food frequency questionnaire was given to assess habitual consumption of foods and beverages in the year preceding the start of the study (24, 25). The assessment of fruits and vegetables was described in detail previously (26) and covered all types of fruits and vegetables regularly eaten in 1986 with the exception of chicory, red cabbage, cucumber, and broccoli; broccoli was hardly sold in the Netherlands at that time. The Dutch Food Composition Table was used to calculate the mean daily intake of vitamin C and dietary fiber (27). Intake of α-carotene, β-carotene, lutein/zeaxanthin, β-cryptoxanthin, and lycopene was calculated as described by Goldbohm et al. (28). Dietary intake of folate was calculated using data from a validated liquid chromatography trienzyme method (29) used to analyze the 125 most important foods contributing to folate intake in the Netherlands (30).

The food frequency questionnaire was validated against a 9-day dietary record, yielding correlation coefficients of 0.60 for fruits, 0.38 for vegetables, 0.74 for dietary fiber, and 0.58 for vitamin C (24). Consumption of fruits was, on average, underestimated (mean daily consumption based on dietary records: 207 g, on questionnaires: 189 g) and consumption of vegetables was overestimated by the questionnaire compared with the dietary record (records: 160 g/d, questionnaires: 189 g/d; ref. 24). The reproducibility of the questionnaire was determined from five annually repeated questionnaire administrations in five independent random samples of 400 cohort members, resulting in a test-retest correlation averaged over all nutrients of 0.66. The average decline in correlation amounted to 0.07 after 5 years (25).

**Study Population.** Twenty-seven colon cancer cases and 215 subcohort members who left ≥60 items blank on the food frequency questionnaires and also reported eating <35 items at least once a month, or who left one or more item blocks (groups of items, e.g., beverages) empty were excluded (24). Because questions on vegetables appeared early in the questionnaire, some participants were more prone to make errors on this item block. When more than three errors were made in the various vegetable questions, the participant was excluded from the analyses on vegetables. This resulted in a final study population of 441 colon cancer cases and 3,048 subcohort members for consumption of fruits and nutrients and 422 colon cancer cases and 2,884 subcohort members for consumption of vegetables.

**Data Analysis.** We examined the distributions of fruits, vegetables, several of their subgroups, related nutrients, and supplements (listed in Table 2). Because 51% of the study population indicated to drink fruit juices other than fresh orange juice less than once per month, these were not examined separately.

Cases and subcohort members were divided into tertiles based on the distribution among subcohort members. First, distributions among patients with hMLH1 protein–deficient and –proficient colon cancer were compared using logistic regression to evaluate etiologic heterogeneity (case-case analyses). Thereafter, Cox proportional hazard models were used to compute incidence rate ratios and 95% confidence intervals for hMLH1 protein–deficient and –proficient colon cancer separately. In the models, the total person-years at risk were estimated from the subcohort (31) and the robust Huber-White sandwich estimator was used to estimate SEs, accounting for the additional variance introduced by sampling from

the cohort. The proportional hazard assumption was evaluated by visual examination of plots of scaled Schoenfeld residuals versus time in combination with a formal test (32, 33) and was met for all presented models.

The standard model included gender, age, family history of colorectal cancer, and total energy intake (as continuous variable). Body mass index, physical activity; consumption of fresh (red) meat, meat products, fish, and alcohol; intake of total fat, calcium, and methionine; use of hormone replacement therapy (set to zero for men and nonusers); and smoking (in separate analyses: current/ever/never, duration of smoking, frequency of smoking, and pack-years) were considered as potential confounders. None of these factors changed the rate ratios with >10% and were therefore not added to the standard model. An ordinal score value based on the median value within each tertile in the subcohort was used to test for trend across tertiles.

Three types of sensitivity analyses were conducted. First, subanalyses on proximal colon cancers were conducted to evaluate whether differential associations of fruits and vegetables with hMLH1 protein–deficient and –proficient colon cancer could be attributed to differences in tumor location. Second, we checked whether associations remained similar when restricting to individuals without a positive family history of colorectal cancer in whom hMLH1 germ line mutations are unlikely to occur. Third, we evaluated whether the associations were driven by the higher consumption of fruits and vegetables by women by using gender-specific tertiles. All reported $P$ values are two sided and considered statistically significant if <0.05. The analyses were conducted using Intercooled STATA for Windows 8.0 (Stata Corp., College Station, TX).

## Results

Table 1 describes baseline characteristics of the study population. Cases were older than subcohort members and they, especially those with hMLH1 protein–proficient colon cancers (hMLH1+), more often had a positive family history of colorectal cancer. In comparison with the subcohort, the percentage of men was lower among cases with hMLH1 protein–deficient colon cancer (hMLH1−) and higher among hMLH1+ cases. There were more ex smokers among hMLH1+ cases, but smokers among hMLH1− cases smoked more pack-years than smokers among subcohort members did. Total energy intake, intake of macronutrients and methionine, and alcohol consumption were relatively low among hMLH1− cases, whereas the corresponding distributions among hMLH1+ cases and subcohort members were similar. In particular, hMLH1− cases tended to consume more meat and less fish than the subcohort did. hMLH1− tumors were more often proximally located, showed a less advanced stage, and were more frequently poorly or not differentiated compared with hMLH1+ tumors.

Table 2 presents consumption of fruits and vegetables, related micronutrients, and supplements in the study population. Fruit consumption was lower among hMLH1− colon cancer cases than among hMLH1+ cases and subcohort members. The same applies to vitamin C, β-cryptoxanthin, folate, and dietary fiber, albeit less strikingly. Consumption of vegetables and intake of α-carotene and lycopene was somewhat higher among hMLH1− cases and somewhat lower among hMLH1+ cases when compared with the subcohort. The same was true for most subgroups of fruits and vegetables.

**Table 1. Baseline demographic, lifestyle, dietary, and tumor characteristics of patients with hMLH1 protein deficient (hMLH1−) and proficient (hMLH1+) colon cancer and subcohort members**

| Characteristic | Colon cancer cases | | Subcohort ($n$ = 3,048) |
|---|---|---|---|
| | hMLH1− ($n$ = 54) | hMLH1+ ($n$ = 387) | |
| Demographic and lifestyle factors | | | |
| Age (y), mean ± SD | 62.8 ± 4.5 | 63.0 ± 4.0 | 61.4 ± 4.2 |
| Gender, $n$ (% men) | 22 (40.7) | 216 (55.8) | 1,475 (48.4) |
| Family history of colorectal cancer, $n$ (%) | 5 (9.3) | 49 (12.7) | 170 (5.6) |
| Body mass index (kg/m$^2$), mean ± SD* | 25.6 ± 3.5 | 25.6 ± 3.2 | 25.1 ± 3.1 |
| Physical activity, >90 min/d, $n$ (%)† | 20 (37.0) | 92 (24.2) | 806 (26.8) |
| Cigarette smoking status | | | |
| Never, $n$ (%) | 18 (33.0) | 144 (37.2) | 1,128 (37.0) |
| Ex smoker, $n$ (%) | 20 (37.0) | 174 (45.0) | 1,074 (35.2) |
| Smoker, $n$ (%) | 16 (29.6) | 69 (17.8) | 846 (27.8) |
| Pack-years among (ex) smokers, median (p25; p75)‡ | 26.7 (10.9; 32.4) | 20.0 (7.8; 35.3) | 18.2 (8.5; 32.0) |
| Alcohol (g/d), median (p25; p75)§ | 1.8 (0.00; 15.0) | 4.0 (0.00; 18.0) | 4.2 (0.25; 14.5) |
| Dietary characteristics‖ | | | |
| Total energy intake (kJ/d), mean ± SD | 7,505 ± 1,718 | 8,056 ± 2,063 | 8,029 ± 2,168 |
| Protein (g/d), mean ± SD | 66.4 ± 13.8 | 70.9 ± 16.4 | 70.6 ± 17.1 |
| Carbohydrates (g/d), median (p25; p75) | 176.8 (150.1; 206.9) | 191.7 (160.5; 239.5) | 193.6 (158.2; 236.2) |
| Fat (g/d), median (p25; p75) | 79.7 ± 24.5 | 84.0 ± 26.3 | 83.9 ± 27.8 |
| Fresh meat (g/d), median (p25; p75) | 93.2 (77.5; 121.6) | 99.3 (75.7; 121.8) | 97.8 (74.5; 123.8) |
| Fish and shellfish (g/d), median (p25; p75) | 6.2 (0.0; 13.2) | 6.9 (0.0; 19.9) | 8.5 (0.0; 19.9) |
| Calcium (mg/d), median (p25; p75) | 865.6 (657.0; 1,036) | 908.8 (694.0; 1,107) | 888.6 (700.0; 1,106) |
| Methionine (mg/d), mean ± SD | 1,520 ± 337 | 1,607 ± 383 | 1,599 ± 406 |
| Tumor-related characteristics | | | |
| Proximally located, $n$ (%)¶ | 48 (88.9) | 171 (45.4) | — |
| Dukes' stages C or D, $n$ (%)** | 14 (26.9) | 158 (44.3) | — |
| Poorly or undifferentiated tumor, $n$ (%)†† | 18 (40.9) | 60 (17.5) | — |
| Age at diagnosis, mean ± SD | 68.6 ± 4.8 | 68.5 ± 4.1 | — |

*For body mass index, figures are based upon $n$ = 54 for hMLH1−, $n$ = 374 for hMLH1+, and $n$ = 2,948 subcohort members.
†For physical activity, figures are based upon $n$ = 54 for hMLH1−, $n$ = 381 for hMLH1+, and $n$ = 3,007 subcohort members.
‡For pack-years, figures are based upon $n$ = 33 for hMLH1−, $n$ = 227 for hMLH1+, and $n$ = 1,779 subcohort members.
§For alcohol, figures are based upon $n$ = 54 for hMLH1−, $n$ = 383 for hMLH1+, and $n$ = 2,957 subcohort members.
‖Other than fruits and vegetables and related nutrients.
¶For 10 hMLH1+ colon cancers, the exact location is unknown.
**Dukes' stage was known for 52 hMLH1− and 357 hMLH1− colon cancers.
††Differentiation grade is known for 44 hMLH1− and 342 hMLH1+ colon cancers.

**Table 2. Consumption of fruits, vegetables, related micronutrients, and supplements among patients with hMLH1 protein deficient (hMLH1−) and proficient (hMLH1+) colon cancer and subcohort members**

| Characteristic | Colon cancer cases | | Subcohort |
|---|---|---|---|
| | hMLH1− | hMLH1+ | |
| Fruits, median (p25; p75) | *n* = 54 | *n* = 387 | *n* = 3,048 |
| Total fruit consumption (g/d)*,† | 120.3 (69.6; 203.0) | 159.4 (89.3; 237.4) | 156.8 (95.1; 235.1) |
| Citrus fruits and fresh citrus juices (g/d) | 59.5 (11.4; 91.5) | 64.2 (21.4; 107.7) | 59.7 (20.7; 110.5) |
| Apples and pears (g/d)‡ | 51.3 (17.1; 97.2) | 80.1 (25.6; 115.8) | 80.1 (25.0; 119.7) |
| Other fruits (g/d) | 14.5 (6.0; 26.2) | 14.3 (6.3; 29.6) | 15.6 (6.5; 29.4) |
| Vegetables, median (p25; p75) | *n* = 54 | *n* = 368 | *n* = 2,884 |
| Total vegetable consumption (g/d)§ | 181.0 (133.8; 237.6) | 173.2 (130.0; 227.8) | 177.9 (136.8; 227.9) |
| Cruciferous vegetables (g/d)‖ | 29.0 (19.9; 38.4) | 27.2 (17.7; 41.6) | 28.3 (18.3; 42.0) |
| Green leafy vegetables, raw (g/d) | 7.1 (4.4; 10.7) | 7.1 (3.6; 13.6) | 7.1 (3.6; 14.1) |
| Green leafy vegetables, cooked (g/d) | 19.7 (10.2; 32.2) | 18.2 (9.1; 26.7) | 19.0 (10.2; 29.1) |
| Allium vegetables (g/d) | 30.9 (13.5; 41.6) | 21.9 (11.0; 43.0) | 24.0 (11.0; 40.9) |
| Carrots (g/d) | 7.5 (2.7; 16.7) | 8.1 (3.5; 14.1) | 8.6 (3.7; 15.5) |
| Tomatoes (g/d) | 18.8 (9.4; 37.6) | 18.8 (4.7; 32.9) | 18.8 (9.4; 32.9) |
| Legumes (g/d)¶ | 24.0 (13.8; 37.3) | 28.5 (17.0; 42.9) | 28.2 (17.6; 41.6) |
| Other vegetables (g/d)‖ | 20.2 (12.9; 37.0) | 22.0 (12.7; 31.2) | 21.2 (13.3; 31.3) |
| Micronutrients**, median (p25; p75) | *n* = 54 | *n* = 387 | *n* = 3,048 |
| α-carotene (μg/d) | 587.4 (295.3; 921.9) | 567.4 (348.7; 911.6) | 565.1 (333.8; 897.6) |
| β-carotene (μg/d) | 2,645 (1,880; 3,313) | 2,706 (1,981; 3,565) | 2,653 (1,944; 3,578) |
| Vitamin C (mg/d) | 94.1 (66.2; 116.6) | 96.1 (76.5; 127.7) | 97.0 (72.4; 127.1) |
| Lutein/zeaxanthin (μg/d) | 2,323 (1,756; 2,863) | 2,237 (1,738; 2,955) | 2,341 (1,770; 3,001) |
| β-cryptoxanthin (μg/d) | 122.7 (30.6; 255.8) | 132.1 (49.5; 270.5) | 128.8 (47.7; 265.9) |
| Lycopene (μg/d) | 865.5 (383.4; 1,399) | 798.0 (382.5; 1,336) | 808.8 (406.2; 1,332) |
| Folate (μg/d) | 190.7 (158.5; 234.1) | 197.3 (162.0; 247.3) | 200.7 (165.6; 243.9) |
| Total dietary fiber (g/d) | 25.4 (21.7; 28.7) | 26.1 (21.2; 32.5) | 26.1 (21.4; 31.5) |
| Users of supplements, *n* (%) | *n* = 54 | *n* = 387 | *n* = 3,048 |
| Multivitamins or minerals | 0 (0.0) | 15 (3.9) | 145 (4.8) |
| Vitamin C | 3 (5.6) | 21 (5.4) | 187 (6.1) |
| Garlic supplement | 7 (13.0) | 34 (8.8) | 271 (8.9) |

*Includes fruits noted in an open-ended question on frequently consumed items not listed in the questionnaire.
†Processed citrus fruit juices are not included in total fruits.
‡Includes apple sauce.
§Excludes potatoes but includes vegetables noted in an open-ended question on frequently consumed items not listed in the questionnaire.
‖Sauerkraut is not included in the group of cruciferous, but of other vegetables, because a lot of potential anticancer agents are destroyed during its processing.
¶Also includes dried pulses.
**Related to fruits and vegetables.

In Table 3, odds ratios and 95% confidence intervals are presented for hMLH1− and hMLH1+ colon cancer versus each other, as well incidence rate ratios and corresponding 95% confidence intervals for hMLH1− and hMLH1+ colon cancer calculated by using the subcohort. Consumption of fruits was differentially associated with hMLH1− and hMLH1+ colon cancer; their consumption was inversely associated with hMLH1− colon cancer but not with hMLH1+ colon cancer. The same pattern was observed in the subgroups citrus fruits and apples/pears, although statistical significance was not reached. Consumption of other types of fruits as well as vegetables was not associated with hMLH1− and hMLH1+ colon cancer. No associations were found with α-carotene, β-carotene, vitamin C, lutein/zeaxanthin, β-cryptoxanthin, lycopene, folate, and dietary fiber (results only shown for folate and vitamin C), nor did addition of these nutrients to the models on fruits and vegetables alter the rate ratios by >10%.

Sensitivity analyses are presented in Table 4. The estimates pointed in the same direction as the main estimates, although the strength of the association between fruits and hMLH1 protein–deficient colon cancer decreased when we applied gender-specific tertiles.

## Discussion

We observed that fruit consumption was inversely associated with hMLH1 protein–deficient colon cancer but not with hMLH1 protein–proficient colon cancer. This observation is in line with inverse associations between consumption of citrus fruits and between consumption of apples/pears and hMLH1 protein–deficient colon cancer, although these associations were not statistically significant. No association between other types of fruits and hMLH1 protein–deficient colon cancer was found, and the association of fruit consumption with hMLH1 protein–deficient cancer could not be attributed to antioxidants related to fruits and vegetables, folate, and dietary fiber. Consumption of vegetables was not associated with either hMLH1 protein–deficient and –proficient colon cancer.

Our results on fruits are in line with two Dutch case-control studies. In the first, fruit consumption was associated with MSI-high colon cancer with hypermethylated hMLH1 but not with MSI-high colon cancer without hypermethylated hMLH1 (16). In the other, fruit consumption was associated with a lower risk of colorectal cancer among hereditary nonpolyposis colorectal cancer (HNPCC) patients (34). Conversely, in an American case-control study, no associations between fruit consumption and colon cancer with or without MSI were observed (17). With regard to consumption of vegetables, in all studies conducted thus far including the current study, case-case analyses showed no differential association for MMR-proficient and -deficient colon cancer (16, 17). No evidence was found for an association with vegetables for hMLH1 protein–deficient and –proficient colon cancer in our case-cohort analyses and the study on HNPCC-related tumors (34), whereas consumption of vegetables was inversely associated with sporadic colon cancer with and without MSI in two other studies (16, 17). Intake of antioxidants related to fruits and vegetables, dietary fiber, and folate was not associated with either type of tumor in this

**Table 3. Odds ratios, incidence rate ratios and 95% confidence intervals for hMLH1 protein deficient (hMLH1−) and proficient (hMLH1+) colon cancer according to consumption categories of fruits, vegetables, folate, and vitamin C**

| Subgroup | Consumption level | Colon cancer cases (*n*) | | Person-years of observation in subcohort | Case-case analyses* | Case-cohort analyses† | |
|---|---|---|---|---|---|---|---|
| | | hMLH1− | hMLH1+ | | hMLH1− versus hMLH1+ | hMLH1− | hMLH1+ |
| All fruits | 0 to <116.0 g/d | 26 | 130 | 4,920 | 1 (reference) | 1 (reference) | 1 (reference) |
| | 116.0 to <204.6 g/d | 15 | 125 | 4,898 | 0.55 (0.27-1.11) | 0.55 (0.29-1.05) | 0.97 (0.74-1.26) |
| | ≥204.6 g/d | 13 | 132 | 4,942 | 0.48 (0.23-0.98) | 0.46 (0.23-0.90) | 1.03 (0.78-1.35) |
| $P_{trend}$ | | | | | 0.046 | 0.029 | 0.81 |
| Citrus and fresh citrus juices | 0 to <31.8 g/d | 21 | 121 | 4,922 | 1 (reference) | 1 (reference) | 1 (reference) |
| | 31.8 to <89.8 g/d | 18 | 141 | 4,884 | 0.67 (0.34-1.32) | 0.81 (0.43-1.52) | 1.16 (0.89-1.51) |
| | ≥89.8 g/d | 15 | 125 | 4,954 | 0.66 (0.32-1.36) | 0.67 (0.34-1.32) | 1.10 (0.83-1.45) |
| $P_{trend}$ | | | | | 0.25 | 0.26 | 0.60 |
| Apples/pears | 0 to <44.5 g/d | 21 | 117 | 4,477 | 1 (reference) | 1 (reference) | 1 (reference) |
| | 44.5 to <114.3 g/d | 20 | 137 | 5,281 | 0.80 (0.41-1.55) | 0.79 (0.42-1.48) | 1.02 (0.78-1.33) |
| | ≥114.3 g/d | 13 | 133 | 5,002 | 0.53 (0.25-1.11) | 0.55 (0.27-1.09) | 1.04 (0.79-1.37) |
| $P_{trend}$ | | | | | 0.094 | 0.085 | 0.76 |
| Other | 0 to <8.9 g/d | 20 | 142 | 4,903 | 1 (reference) | 1 (reference) | 1 (reference) |
| | 8.9 to <23.5 g/d | 19 | 119 | 4,903 | 1.22 (0.61-2.43) | 0.96 (0.51-1.81) | 0.83 (0.64-1.08) |
| | ≥23.5 g/d | 15 | 126 | 4,955 | 0.90 (0.43-1.86) | 0.78 (0.39-1.56) | 0.89 (0.68-1.16) |
| $P_{trend}$ | | | | | 0.68 | 0.46 | 0.53 |
| All vegetables | 0 to <150.6 g/d | 21 | 137 | 4,630 | 1 (reference) | 1 (reference) | 1 (reference) |
| | 150.6 to <209.5 g/d | 16 | 108 | 4,659 | 1.07 (0.52-2.17) | 0.80 (0.41-1.55) | 0.81 (0.62-1.07) |
| | ≥209.5 g/d | 17 | 123 | 4,687 | 1.00 (0.49-2.01) | 0.86 (0.45-1.65) | 0.94 (0.72-1.23) |
| $P_{trend}$ | | | | | 0.99 | 0.67 | 0.72 |
| Micronutrients | | | | | | | |
| Folate | <177.2 μg/d | 23 | 136 | 4,917 | 1 (reference) | 1 (reference) | 1 (reference) |
| | 177.2 to <225.6 μg/d | 15 | 114 | 4,943 | 0.91 (0.44-1.87) | 0.75 (0.38-1.52) | 0.83 (0.62-1.10) |
| | ≥225.6 μg/d | 16 | 137 | 4,901 | 0.93 (0.43-2.05) | 0.92 (0.43-2.00) | 1.04 (0.78-1.39) |
| $P_{trend}$ | | | | | 0.85 | 0.86 | 0.63 |
| Vitamin C | <80.7 mg/d | 18 | 120 | 4,919 | 1 (reference) | 1 (reference) | 1 (reference) |
| | 80.7 to <115.7 mg/d | 22 | 142 | 4,910 | 1.08 (0.55-2.13) | 1.25 (0.65-2.38) | 1.18 (0.90-1.54) |
| | ≥115.7 mg/d | 14 | 125 | 4,932 | 0.83 (0.38-1.79) | 0.81 (0.39-1.68) | 1.08 (0.81-1.43) |
| $P_{trend}$ | | | | | 0.65 | 0.51 | 0.67 |

*Odds ratios are presented for case-case analyses.
†Incidence rate ratios are presented for case-cohort analyses.

study. In contrast, intake of folate, dietary fiber, and β-carotene was inversely associated with colon cancer with and without MSI in an American case-control study (17), although only the associations of folate and fiber with MSI-positive colon cancer were statistically significant.

Hypermethylation status of the promoter region of the *hMLH1* gene determines hMLH1 protein expression in the majority (80-90%) of the sporadic colon cancers (2-8), whereas germ line mutations in the *hMLH1* gene are responsible for lack of expression in 50% of the HNPCC-related colon cancers (35). Our classification of hMLH1 protein–deficient and –proficient colon cancers is likely to be primarily determined by hypermethylation status of the promoter region of the *hMLH1* gene. The youngest patient in our study population was diagnosed at age 57; the proportion of HNPCC was only around 2% in other unselected series (8, 36), and the parameter estimates pointed in the same direction when we restricted the analyses to individuals without a family history of colorectal cancer.

Because dietary methyl donors such as folate are thought to affect methylation status of cancer genes (37, 38), the association between fruits and hMLH1 protein– deficient colon cancer may arise because fruits affect hypermethylation of the promoter region of the *hMLH1* gene. We did not observe an association between intake of folate and the rate of hMLH1 protein–deficient colon cancer. Thus, folate, and its potential

**Table 4. Sensitivity analyses: odds ratios and incidence rate ratios for hMLH1 protein deficient (hMLH1−) and proficient (hMLH1+) colon cancer comparing the third tertile of consumption of fruits or vegetables with the first**

| Comparison | Case-case analyses* | Case-cohort analyses† | |
|---|---|---|---|
| | hMLH1− versus hMLH1+ | hMLH1− | HMLH1+ |
| Total fruit consumption | | | |
| Estimate in Table 3 | 0.48 | 0.46 | 1.03 |
| Gender-specific tertiles | 0.57 | 0.54 | 1.03 |
| Cases with proximal tumors only | 0.44 | 0.51 | 1.20 |
| Negative family history of colorectal cancer | 0.60 | 0.55 | 1.00 |
| Total vegetables | | | |
| Estimate in Table 3 | 1.00 | 0.86 | 0.94 |
| Gender-specific tertiles | 0.93 | 0.82 | 0.96 |
| Cases with proximal tumors only | 0.70 | 0.70 | 1.07 |
| Negative family history of colorectal cancer | 1.24 | 1.04 | 0.91 |

*Odds ratios are presented for case-case analyses.
†Incidence rate ratios are presented for case-cohort analyses.

role in methylation of hMLH1, by itself seems insufficient to explain the association between fruit consumption and hMLH1 protein–deficient colon cancer. As fruit consumption was also inversely associated with HNPCC-related tumors (34), our findings may suggest that fruits play a role in pathways involved in hMLH1 protein–deficient colon carcinogenesis that are determined by other factors than hypermethylation of the hMLH1 promoter region. However, another study found that an inverse association between fruits and MSI-high colorectal cancer was restricted to tumors with hypermethylated hMLH1, whereas no associations were observed with the methyl donors folate and alcohol (16). Perhaps components in fruits other than folate influence hypermethylation. Pathways involved are unlikely to depend solely on fruit-related antioxidants and dietary fiber, because they were not associated with hMLH1 protein–deficient colon cancer and hardly influenced the association between fruits and hMLH1 protein–deficient colon cancer. Hence, the mechanism behind this association remains to be elucidated.

Our classification of hMLH1 protein–deficient and –proficient tumors seems reliable. First, the percentage of hMLH1 protein–deficient colon cancers in our study corresponds with the literature (10, 11, 16). Second, we found a high agreement with classification based on the MSI marker BAT-26[6] as expected (39-41).

Third, the sensitivity (94%) and specificity (100%) of hMLH1 immunohistochemistry was high in an unselected series of colorectal cancer with MSI testing as golden standard (8). However, a minor percentage of our tumors may show MMR deficiency because of mutations in other MMR genes than hMLH1 and some hMLH1-proficient tumors may have nonfunctional protein, resulting in a bias towards the null.

hMLH1 protein–deficient colon cancers occur mostly in the proximal colon. We observed that the association remained when we restricted the analyses to proximal tumors. Thus, distinction of hMLH1 protein deficient and proficient colon cancer, in addition to tumor location, provides greater information than tumor location by itself.

The strengths of this study include its large size, prospective nature with 7.3 years of follow-up, the fact that it is based on an unselected series of colon cancer, a very low chance of selective dropout because no subcohort members were lost to follow-up, and high completeness of follow-up for cancer incidence.

As in other epidemiologic studies, it is impossible to completely rule out bias that occurs when factors determining tumor block availability or successful hMLH1 immunohistochemistry are associated with fruit or vegetable consumption and hMLH1 expression status (42, 43). It is unlikely that such a bias has affected our results considerably. We did not observe differences in distribution of Dukes' stages and differentiation grade between tumors of patients in which hMLH1 expression status could be determined and tumors in which we could not assess this status. Moreover, our study population has, on average, the same age and contains a similar percentage of individuals with a positive family history as the complete study population after 7.3 years of follow-up.

Because diet was assessed before the occurrence of disease and the initial 2.3 years of follow-up were excluded, changes in diet due to (subclinical) disease occurrence cannot have affected the assessment. It is no problem that we assessed diet only once, because the small decline in test-retest correlation coefficients over time (0.07 over 5 years) in the reproducibility study of our food frequency questionnaire (25) suggests that the exposure is relatively stable over time. However, assessment of diet brings along systematic and random errors. The correlation coefficients between fruit and vegetable consumption assessed by our food frequency questionnaire and by 9-day dietary records are comparable with those of questionnaires used in other prospective cohort studies (44, 45), but the moderate ability of the questionnaire to rank according to vegetable consumption ($r = 0.38$ between assessment according to questionnaire and dietary records; ref. 24) results in an attenuation of the associations. Given the substantial correlation coefficient between assessment of fruits by questionnaires and dietary records ($r = 0.60$; ref. 24), it is less likely that misclassification has affected our final conclusion for fruits.

After 7.3 years of follow-up in the Netherlands Cohort Study on diet and cancer, we observed an inverse association between fruits and colon cancer, confined to the subgroup of hMLH1 protein–deficient colon cancer cases. Our findings may provide clues for elucidating pathways involved in carcinogenesis of MMR-deficient colon cancer.

[6] M. Lüchtenborg, et al. Mutations in APC, CTNNB1 and K-ras genes and expression of hMLH1 in sporadic colorectal cancer, submitted for publication.

## Acknowledgments

We thank the participants of our study; Dr. M. Brink for collecting tumor specimen; H. Braam and G. Roemen for laboratory assistance; Prof. Dr. J.H.J.M. van Krieken for reviewing hMLH1 scoring; Drs. A. Volovics and A. Kester for statistical advice; Drs. L. Schouten, A. de Bruïne, and L.E. Voorrips, S. van de Crommert, H. Brants, J. Nelissen, C. de Zwart, M. Moll, W. van Dijk, M. Jansen, and A. Pisters, for assistance; H. van Montfort, T. van Moergastel, L. van den Bosch, and R. Schmeitz for programming assistance; Prof. Dr. J.W. Arends for his participation in the initiation of this study; the regional cancer registries (IKA, IKL, IKMN, IKN, IKO, IKR, IKST, IKW, IKZ); the Dutch nationwide network; registry of histopathology and cytopathology (PALGA); and the departments of pathology of the following hospitals for providing the tissue blocks: Academisch Ziekenhuis Nijmegen Sint Radboud, Academisch Ziekenhuis Groningen, Rijnland Ziekenhuis, Antoni van Leeuwenhoek Ziekenhuis, Academisch Ziekenhuis Rotterdam, Stichting Laboratorium Pathologie Oost Nederland, Pathologisch Instituut Utrecht, Ziekenhuis Rijnstate Arnhem, Laboratorium Volksgezondheid Leeuwarden, Ziekenhuis Bethesda, Stichting Samenwerkende Ziekenhuizen Oost Groningen, Martini Ziekenhuis Groningen, Stichting Samenwerkende Delftse Ziekenhuizen, Leyenburg Ziekenhuis, Academisch Ziekenhuis Vrije Universiteit, Academisch Medisch Centrum, Sint Franciscus Ziekenhuis, Dr. Daniel den Hoed Kliniek, Academisch Ziekenhuis Maastricht, Stichting Laboratorium Goudse Ziekenhuizen, Canisius-Wilhelmina Ziekenhuis, Slotervaart Ziekenhuis, Maaslandziekenhuis, Atrium Heerlen, Atrium Kerkrade and Brunssum, Microbiologie St Medische Stedendriehoek, IJsselmeer Ziekenhuizen, Ziekenhuis Centrum Apeldoorn, Isala Klinieken, Elkerliek ziekenhuis, Groot Ziekengasthuis, Ziekenhuis Gooi-Noord, Medisch Centrum Alkmaar, Regionaal Pathologisch en Cytologisch Laboratorium voor Eemland en Noord-West Veluwe, Diakonesse Ziekenhuis, Sint Antonius Ziekenhuis, Onze Lieve Vrouwe Gasthuis, St. Lucas Andreas Ziekenhuis, Pathologisch Anatomisch Laboratorium SPALK, Ziekenhuis De Heel, Diakonessenhuis, Rode Kruis Ziekenhuis, Ziekenhuis Bronovo, Laurentius Ziekenhuis Roermond, Pathologisch Anatomisch Laboratorium Dordrecht, Zuiderziekenhuis, Sint Clara Ziekenhuis, Medisch Centrum Haaglanden, Stichting Streeklaboratorium Zeeland, Sint Elisabeth Ziekenhuis, Catharina-ziekenhuis, Sint Maartensgasthuis, and Spaarne Ziekenhuis.

## References

1. Atkin NB. Microsatellite instability. Cytogenet Cell Genet 2001;92:177–81.
2. Herman JG, Umar A, Polyak K, et al. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci U S A 1998;95:6870–5.
3. Wheeler JM, Loukola A, Aaltonen LA, Mortensen NJ, Bodmer WF. The role of hypermethylation of the hMLH1 promoter region in HNPCC versus MSI+ sporadic colorectal cancers. J Med Genet 2000;37:588–92.
4. Kim HC, Kim CN, Yu CS, Roh SA, Kim JC. Methylation of the hMLH1 and hMSH2 promoter in early-onset sporadic colorectal carcinomas with microsatellite instability. Int J Colorectal Dis 2003;18:196–202.
5. Kuismanen SA, Holmberg MT, Salovaara R, De la Chapelle A, Peltomaki P. Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers. Am J Pathol 2000;156:1773–9.

6. Kamory E, Kolacsek O, Otto S, Csuka O. hMLH1 and hMSH2 somatic inactivation mechanisms in sporadic colorectal cancer patients. Pathol Oncol Res 2003;9:236–41.
7. Cunningham JM, Christensen ER, Tester DJ, et al. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. Cancer Res 1998;58:3455–60.
8. Cunningham JM, Kim CY, Christensen ER, et al. The frequency of hereditary defective mismatch repair in a prospective series of unselected colorectal carcinomas. Am J Hum Genet 2001;69:780–90.
9. Chao A, Gilliland F, Willman C, et al. Patient and tumor characteristics of colon cancers with microsatellite instability: a population-based study. Cancer Epidemiol Biomarkers Prev 2000;9:539–44.
10. Ward R, Meagher A, Tomlinson I, et al. Microsatellite instability and the clinicopathological features of sporadic colorectal cancer. Gut 2001;48: 821–9.
11. Miyakura Y, Sugano K, Konishi F, et al. Extensive methylation of hMLH1 promoter region predominates in proximal colon cancer with microsatellite instability. Gastroenterology 2001;121:1300–9.
12. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. Science 1993;260:816–9.
13. Lothe RA, Peltomaki P, Meling GI, et al. Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history. Cancer Res 1993;53:5849–52.
14. Davis CD, Uthus EO. DNA methylation, cancer susceptibility, and nutrient interactions. Exp Biol Med (Maywood) 2004;229:988–95.
15. Glaab WE, Hill RB, Skopek TR. Suppression of spontaneous and hydrogen peroxide-induced mutagenesis by the antioxidant ascorbate in mismatch repair-deficient human colon cancer cells. Carcinogenesis 2001;22:1709–13.
16. Diergaarde B, Braam H, Van Muijen GN, Ligtenberg MJ, Kok FJ, Kampman E. Dietary factors and microsatellite instability in sporadic colon carcinomas. Cancer Epidemiol Biomarkers Prev 2003;12:1130–6.
17. Slattery ML, Anderson K, Curtin K, Ma KN, Schaffer D, Samowitz W. Dietary intake and microsatellite instability in colon tumors. Int J Cancer 2001;93:601–7.
18. Van den Brandt PA, Goldbohm RA, Van 't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol 1990;43:285–95.
19. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol 1990;19:553–8.
20. Schouten LJ, Straatman H, Kiemeney LA, Gimbrère CH, Verbeek AL. The capture-recapture method for estimation of cancer registry completeness: a useful tool? Int J Epidemiol 1994;23:1111–6.
21. Iacopetta B. Are there two sides to colorectal cancer? Int J Cancer 2002;101: 403–8.
22. Kapiteijn E, Liefers GJ, Los LC, et al. Mechanisms of oncogenesis in colon versus rectal cancer. J Pathol 2001;195:171–8.
23. Wei EK, Giovannucci E, Wu K, et al. Comparison of risk factors for colon and rectal cancer. Int J Cancer 2004;108:433–42.
24. Goldbohm RA, Van den Brandt PA, Brants HA, et al. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur J Clin Nutr 1994;48:253–65.
25. Goldbohm RA, Van 't Veer P, Van den Brandt PA, et al. Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. Eur J Clin Nutr 1995;49:420–9.
26. Zeegers MP, Goldbohm RA, van den Brandt PA. Consumption of vegetables and fruits and urothelial cancer incidence: a prospective study. Cancer Epidemiol Biomarkers Prev 2001;10:1121–8.
27. Stichting Nederlands Voedingsstoffenbestand. Dutch Food Composition Table 1986-1987; Nederlands voedingsstoffenbestand 1986-1987: The Hague: Voorlichtingsbureau voor de Voeding; 1996.
28. Goldbohm RA, Brants HA, Hulshof KF, Van den Brandt PA. The contribution of various foods to intake of vitamin A and carotenoids in the Netherlands. Int J Vitam Nutr Res 1998;68:378–83.
29. Konings EJ. A validated liquid chromatographic method for determining folates in vegetables, milk powder, liver, and flour. J AOAC Int 1999;82:119–27.
30. Konings EJ, Roomans HH, Dorant E, Goldbohm RA, Saris WH, Van den Brandt PA. Folate intake of the Dutch population according to newly established liquid chromatography data for foods. Am J Clin Nutr 2001;73: 765–76.
31. Volovics A, Van den Brandt PA. Methods for the analyses of case-cohort studies. Biomed J 1997;2:195–214.
32. Schoenfeld DA. Partial residuals for the proportional hazards regression model. Biometrika 1982;69:239–41.
33. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 1994;81:515–26.
34. Diergaarde B, Braam H, Vasen HF, et al. Dietary factors and HNPCC-associated colorectal tumors. Third Annual AACR International Conference Frontiers in Cancer Prevention Research. Oct 16-20; Seattle, WA, USA, 2004:135.
35. Peltomaki P. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Hum Mol Genet 2001;10:735–40.
36. Aaltonen LA, Salovaara R, Kristo P, et al. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. N Engl J Med 1998;338:1481–7.
37. Friso S, Choi SW. Gene-nutrient interactions and DNA methylation. J Nutr 2002;132:2382–7S.
38. Van Engeland M, Weijenberg MP, Roemen GM, et al. Effects of dietary folate and alcohol intake on promoter methylation in sporadic colorectal cancer: the Netherlands cohort study on diet and cancer. Cancer Res 2003;63:3133–7.
39. Cawkwell L, Gray S, Murgatroyd H, et al. Choice of management strategy for colorectal cancer based on a diagnostic immunohistochemical test for defective mismatch repair. Gut 1999;45:409–15.
40. Lindor NM, Burgart LJ, Leontovich O, et al. Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. J Clin Oncol 2002;20:1043–8.
41. Ward R, Meldrum C, Williams R, et al. Impact of microsatellite testing and mismatch repair protein expression on the clinical interpretation of genetic testing in hereditary non-polyposis colorectal cancer. J Cancer Res Clin Oncol 2002;128:403–11.
42. Hoppin JA, Tolbert PE, Taylor JA, Schroeder JC, Holly EA. Potential for selection bias with tumor tissue retrieval in molecular epidemiology studies. Ann Epidemiol 2002;12:1–6.
43. Schroeder JC, Weinberg CR. Use of missing-data methods to correct bias and improve precision in case-control studies in which cases are subtyped but subtype information is incomplete. Am J Epidemiol 2001;154:954–62.
44. Hankin JH, Wilkens LR, Kolonel LN, Yoshizawa CN. Validation of a quantitative diet history method in Hawaii. Am J Epidemiol 1991;133:616–28.
45. Ocké MC, Bueno de Mesquita HB, Goddijn HE, et al. The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. Int J Epidemiol 1997;26 Suppl 1:S37–48.